Với sách anh tiệp

|  |  |
| --- | --- |
| PerCeptron | logistic |
| * Xây dựng hàm loss này dựa trên ý tưởng tìm số lượng điểm bị phân lớp lỗi và tìm cách tối ưu hàm này * Với 1 điểm bị phân lớp lỗi tức sgn(WX) # yi, vậy lên yi \* sgn(WX) = -1 * Vậy loss : * Nhưng do đạo hàm này bị gián đoạn lên họ đã tìm ra loss thay thê là :      * Tối ưu các thứ thì tìm được công thức update W = W + yixi ( với xi là điểm nhãn phân loại lỗi ) | * Xây dựng hàm loss dựa trên likelihood * Và sau đó tìm được công thức update W cho logistic như này |

Với các sách khác thì

Họ dùng MSE làm loss cho cả perceptron và logistic

Lên loss sẽ có dạng :

Perceptron : loss = (y – sgn(WX))^2 (1)

Logistic : loss = (y – sigmoid(WX))^2 (2)

Với MSE cho perceptron : luôn có W = W – learning rate \*( f(x)) , f(x) ở đây là hàm loss

* W = W -  **α\*f’(x) mà f(x) = ½ \* (y – WX)^2 => f’(x) = X\*(yj – dj) với dj là nhãn (label ) còn yi là predict**
* W = W -  **α\*(yj – dj)X**

NOTE : ở đây ko viết đạo hàm sgn mà dùng luôn WX sách anh tiệp đã giải thích còn muốn biết nếu với bài toán đạo hàm thì xem

MSE của logistic

Với MSE logistic : có công thức update W của logistic MSE

W = W – (1)’

Tính đạo hàm của (y – sigmoid(WX))^2 đạo hàm hàm hợp các thứ blab la

Ta có đầu ra dự đoán là yi = sigmoid (WX) (predict )

Đạo hàm hàm sigmoid vd x sẽ ra x\*(1-x)

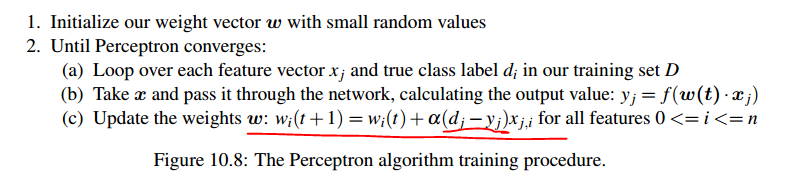
Lên có đạo hàm hàm loss (1)’ = X \* error \* đạo hàm sigmoid ( predict )

⬄ (1)’ = X \* (yi – di) \* (yi \* (1-yi))

Vậy W = W – **α \* (1)’**

**Vậy tóm lại biểu thức MSE cho perceptron có (predict – label) nhìn giống vs cái của logistic nhưng bản chất hình thành lại là 2 nội dung hoàn toàn khác nhau**

**Perceptron MSE :**

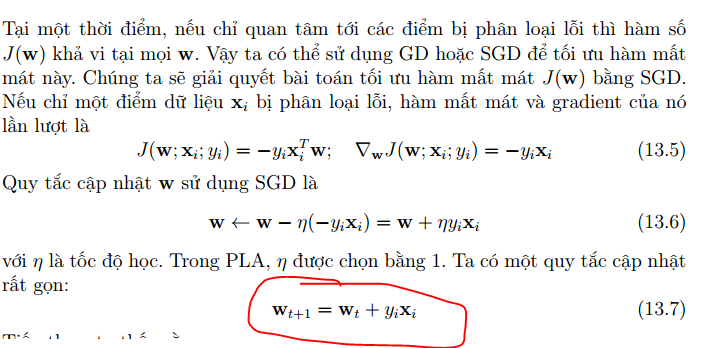
****

**Logistic likelihood :**

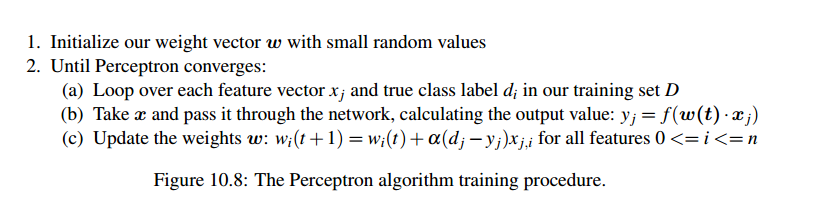
****

**Lưu ý nữa là cách xây dựng update W trong sách a tiệp là dựa vào hàm dấu và các thứ kiến thức liên quan lên có W = W + yixi chứ ko phải dùng MSE như sách khác lên công thức cũng sẽ khác và 3 công thức này hoàn toàn khác nhau vì dùng các loss khác nhau**

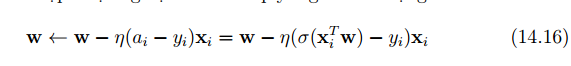
1. **Perceptron loss trên lý thuyết**

****

1. **Perceptron dùng MSE**

****

1. **Logistic likelihood**

****